CSE
COMPUTER SCIENCE & ENGINEERING

SKILL WORKBOOK

17CS2212 MACHINE LEARNING

Team ML
K L UNIVERSITY | MACHINE LEARNING – 17CS3166

SKILL WORKBOOK

STUDENT NAME	
REG. NO	
YEAR	
SEMESTER	
SECTION	
FACULTY	

Table of Contents

Koneru Lakshmaiah Education foundation
(KL DEEMED TO BE UNIVERSITY)

17CS3166 – Skill – Machine Learning

REGISTER NO : NAME: YEAR / SEM / SECTION

S.No	Date	Aim	Observation (10M)	Logic (5M)	Execution (10M)	Result (10 M)	Analysis (5 M)	Viva Voce (10M)	Total Marks (50M)	Signature of Faculty with date
1		Creation of data set using Pandas								
2		Data Manipulation using Numpy								
3		Constructing a linear regression model								

S.No	Date	Aim	Observation (10M)	Logic (5M)	Execution (10M)	Result (10 M)	Analysis (5 M)	Viva Voce (10M)	Total Marks (50M)	Signature of Faculty with date
4		Constructing Decision Tree model								
5		Naïve Bayes Classification								
6		Basic operations on TensorFlow								
7		Regression using TensorFlow								
8		Logistic Regression using TensorFlow								

S.No	Date	Aim	Observation (10M)	Logic (5M)	Execution (10M)	Result (10 M)	Analysis (5 M)	Viva Voce (10M)	Total Marks (50M)	Signature of Faculty with date
9		K – Means Clustering using TensorfFow								
10		Basic Operations using Opencv								
11		Image Processing using Opencv								
12		Video capturing using Opencv								

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SUBJECT CODE: 17CS2212
MACHINE LEARNING

Exercise #1: Creation of data set using Pandas

Date of the Session:___/___/___
Time of the Session:_____to_____

Based on the following queries write a python code in jupyter note book

A) INTRODUCTION TO PANDAS

1. Use series method to convert the following list[‘a’,’b’,’c’,’d’]
2. Given the lists which contain names of students and their scored marks you have to create data frames which club them all using pandas. student list is ['praharsha' ,'mythresh', 'kowshik' ,'pranay', 'srujan' 'raju'] and you are given a 2d matrix which contain their marks. Their marks are [[77,89,92,95], [85,94,97,76], [77,88,99,74], [90,89,92,78], [83,97,80,98],[90,90,80,87]] and subjects list containing ['bio','phy','che', 'mat']
3. Add the sci , total columns to the data frame from which science is the sum of phy, che and bio and total is the sum of all subjects?
4. Display only the marks scored by praharsha and srujan?
5. Display all the students marks who scored 90 + marks in physics?
6. Biology teacher wants to know the total marks achieved by students in biology, Help her by using the sum method in data frame?

B) EXCEL SHEET HANDLING

1. Load the data set and print the data in csv file?
2. Find the sum of all columns in that csv file?
3. Load and just first two columns from all the rows?
4. Plot the graph of the dataset without matplotlib?
5. Find the mean of all the columns?
6. Find the mode, minimum, absolute, standard deviation of the data set?

Writing space for the Problem: (For Student's use only)

Writing space for the Problem: (For Student's use only)

(For Evaluator's use only)

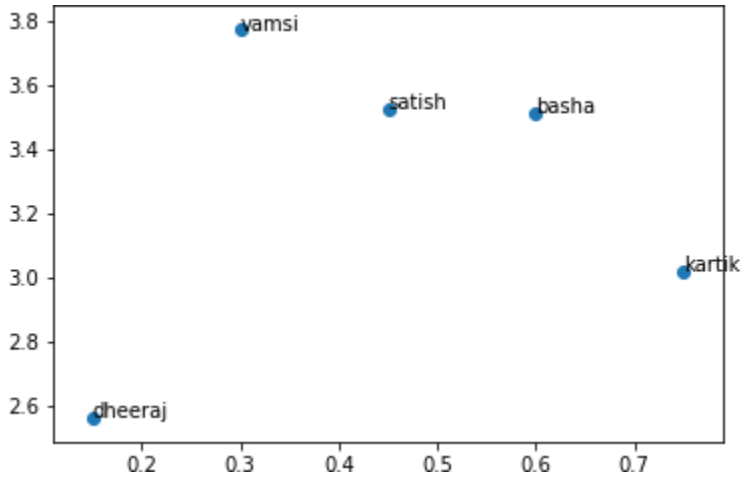
Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #2: Data Manipulation using Numpy

Date of the Session:___/___/___
Time of the Session:_____to_____

1) Write a program of a two-dimensional array using numpy and print elements, reverse the rows of the 2D array and print the elements?
2) Generate the digits up to 15 using the np.arange method?
3) Print the common items from and hu two Numpy arrays and also print the position of the common elements and remove the common items and print the different elements also of two Numpy arrays?
4) Write a numpy program to generate the 10 random integers between 10 and 100.
5) Generate a random 4*4 matrix of integers ranging from 0 to 100 by using numpy and also to use in the nd minimum values?
6) Divide all digits from 1 to 10010011 using linspace method?
7) Find the default plot size in matplotlib plots, print the default size and change the plot size to [10,10] by using the matplotlib library.
8) Create a line plot for the given equation y=x^3 and plot some points using linspace() in the range (-10,10) and also label x-axis, y-axis and give the title of graph cubic function and display the graph
9) Represent a graph by using subplot as two rows and two columns and fill all the positions by plotting y=x^2 and display the graph.
10) Represent two equations $y=x^3$& $y=x^2$ in one graph and use **Legend** function.

11) Generate the graph by using given data as in the given diagram y=[2.56422, 3.77284,3.52623, 3.51468, 3.02199] ,z = [0.15, 0.3, 0.45, 0.6, 0.75] ,n = ['dheeraj', 'vamsi', 'satish', 'basha', 'kartik'].



12) Explain the definition and uses of the following graphs in detail:

1. Factor plot
2. Density plot
3. Box whiskers plot

Y
39343
46205
37731
43525
39891
56642
60150
54445
64445
57189

13) Consider the above data and plot the following graphs and explain them briefly:

i) Bar Plot for analyzing whether '**Y**' values are increasing with '**X**' with width=1/1.5 and in blue colour.

ii) Check the frequency distribution of all '**Y**' values for 20 bins.

iii) Identify the mean and medians of the '**Y**' values by plotting a horizontal notched box-whiskers plot. (The mean should be identified as a point in the graph).

14) Create a random10 X 5 numpy matrix. Convert it into a data frame (Using pandas) with column names ='A', 'B', 'C', 'D' and 'E'. Now, analyze the data frame by drawing an **area plot.**

15) Do the following:

X	Y
4.1	57081
4.5	61111
4.9	67938
5.1	66029
5.3	83088
5.9	81363

i) From the above data, Identify all the points of X, Y as co-ordinates by using Scatter plot.

ii) Create a random numpy array of size 3. Create a data frame named 'Dishes' using the created array with indexes=' Cup Cake', 'Bread' and 'Cookie'. Let us assume the values of the array indicate how much a person likes the respective dishes. Now, draw a **Pie plot** and identify his/her most liked dish.

16) What are the uses of the seaborn library? How is it different from matplotlib?

A	B	C	E
165349.2	136897.8	471784.1	192261.83
162597.7	151377.59	443898.53	191792.06
153441.51	101145.55	407934.54	191050.39
144372.41	118671.85	383199.62	182901.99
142107.34	91391.77	366168.42	166187.94
131876.9	99814.71	362861.36	156991.12
134615.46	147198.87	127716.82	156122.51
130298.13	145530.06	323876.68	155752.6
120542.52	148718.95	311613.29	152211.77
123334.88	108679.17	304981.62	149759.96

17) In the above dataset, Imagine '**E**' is the dependent variable and 'A', 'B', and 'C' are independent. using **Heatmap**, determine which among 'A', 'B' and 'C' has a high impact on 'E'.

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #3: Constructing a linear regression model

Date of the Session:___/___/___
Time of the Session:_____to_____

Build a simple linear regression model using numpy, pandas, matplotlib and write the python code mathematically.

Download the dataset
https://docs.google.com/spreadsheets/d/1CwOTHHvoGSemWSHta5XGvTH-TjUz2gbMv5CtyhUDdUU/edit

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #4: Constructing Decision Tree model

Date of the Session:___/___/___
Time of the Session:_____to_____

You are having a blood donor data taken from the blood transfusion service center now using the data build a decision tree machine learning model to predict the respective person is willing to donate the blood or not.

Download data from

https://docs.google.com/spreadsheets/d/16jxPzQVipo_7rXtR33F7ikmQkP7Cly0VVYVcNpFFDmI/edit#gid=1023336393

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #5: Naïve Bayes Classification

Date of the Session:____/___/___
Time of the Session:_____to_____

We have a sample data set of adult census income across the globe now using the data build a Naïve Bayes machine learning model to predict their class based on their gender, capital gain, capital loss, hours per week and find the accuracy of your model

Download the link from
https://docs.google.com/spreadsheets/d/14Y4YtwTS-vWN04NkvrynDAC9w8DnHGKm4V7YiUA4bss/edit#gid=1395218118

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #6: Basic operations on TensorFlow

Date of the Session:____/___/___
Time of the Session:_____to_____

1) Create two constants and do addition for those constants and print the result using tensorflow?
2) Create two placeholders and do addition for those placeholders and print the result using tensorflow?
3) Create two Variables and do addition for those Variables and print the result using tensorflow?
4) Create one variable, one placeholder, one constant and do multiplication and print the value?
5) Create 4X4 matrix and fill with default value as zeros using tensorflow?
6) Using random function create 4X4 matrix with values range 0<x<1?
7) Do the following:
    i. Create a 4X4 matrix having mean =0 and stddev=1.0 using random() in tensorflow in singleline?
    ii. Create a 4X4 matrix in range minval=0 and maxval=1.0 using random() in tensorflow in singleline?
    iii. Perform multiplication for above two matrices and print then using tensorflow in single line?
8) using linspace command write a code in tensorflow?
9) Create two tensorflow objects and initialize the values at run time and add them and now create a tensor board graph using then in tensorflow?
10) perform a scatter plot using tensorflow with the flowing given data a=[1,2,3,4,5] and b=[3.5,2.2,1,4,9] ?

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #7: Regression using TensorFlow

Date of the Session:____/___/___
Time of the Session:_____to_____

Create a dataset of 100000 using line space and generate noise of size 100000 using np.random.randn

Consider the line y=0.5*x+5

With the x values,generate y values and add noise to the y values.

Create data frame and concatenate it.

Plot the data points.

After predicting the values check the error using reduce.sum.

And optimize the error using gradient descent optimizer.

Create and run the session and plot the respective line.4

take the following values as input for building a model

x=[3.3,4.4,5.5,6.71,6.93,4.168,9.779,6.182,7.59,2.167,
7.042,10.791,5.313,7.997,5.654,9.27,3.1]
y=[1.7,2.76,2.09,3.19,1.694,1.573,3.366,2.596,2.53,1.221,
2.827,3.465,1.65,2.904,2.42,2.94,1.3]

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #8: Logistic Regression using TensorFlow

Date of the Session:____/___/___
Time of the Session:_____to_____

From tensorflow. examples. tutorials. mnist import the input data.

Learning rate=0.01

Training epochs=25

Batch size=100

Display step=1

Use softmax and matmul to predict the output using the matrices.

Calculate the error using reduce_mean.

Optimize the error using gradient descent optimizer.

Create and run the session using logistic regression.

Calculate the accuracy using accuracy.eval.

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #9: K – Means Clustering using TensorFlow

Date of the Session:____/___/___
Time of the Session:_____to_____

From tensorflow. examples. tutorials. mnist import the input data.

num_steps = 50 # Total steps to train

batch_size = 1024 # The number of samples per batch

k = 25 # The number of clusters

num_classes = 10 # The 10 digits

num_features = 784 # Each image is 28x28 pixels

create and run the session.

Using the above features perform K-Means clustering.

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #10: Basic Operations using Opencv

Date of the Session:____/___/____
Time of the Session:_____to_____

1) Write a program which takes a image and give a Title as 'My Image' .Display it in gray scale using opencv?
2) Write a program to reshape the image as 10X8 using opencv?
3) Write a program to convert a image BGR TO RBG and save it using opencv?
4) Write a program to write a name on image and display it using opencv?
5) write a program to read an image from camera and save it as frames when we click space bar and stop it when we click 'ESC' button using opencv?

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #11: Image Processing using Opencv

Date of the Session:___/___/___
Time of the Session:_____to_____

1) load an image and apply some filters named hsv image,hue channel,saturation,value channel and display all of them(Hint:use 0 for hue channel ,1 for saturation and 2 for value channel)

2) load an image and split that into red,blue,green channels and display each channel and merge them again and display the image.

3) load an image with help of open CV and make the following operations

a)make blue and green colour to zero and print the image.

b) make blue and red colour to zero and print the image.

c) make red and green colour to zero and print the image.

4) create a black colour rectangle with open cv and build a circle and a triangle in between the rectangle?

5) create the black colour rectangle with openCV and wite your name in it.

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation:

Exercise #12: Video capturing using Opencv

Date of the Session:___/___/___
Time of the Session:_____to_____

SKILL-12

1) Write a program to run a video which is present in our system in a gray scale using opencv?
2) Write a program to capture a video in HSV, GRAY, RGB using opencv?
3) Write a program to capture a video in HSV, GRAY and save the video in HSV using csv?
4) Write a program to calculate the height and width of a frame and print them and set the height, width of the frame to ([3,3000], [4,3000] using opencv?
5) Do the following using opencv:
    i. Write a program to keep text like width, height on the frame which is visualize from camera?
    ii. Write a program to keep date and time on the frame which is visualize from camera

(For Evaluator's use only)

Comment of the Evaluator (if Any)	Evaluator's Observation Marks Secured:_____out of _____ Full Name of the Evaluator: Signature of the Evaluator Date of Evaluation: